



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/736,273	12/15/2003	Michael Edward Houle	JP920020208US1	3829
36380	7590	05/18/2007		
RICHARD M. GOLDMAN 371 ELAN VILLAGE LANE SUITE 208, CA 95134			EXAMINER LE, MIRANDA	
			ART UNIT 2167	PAPER NUMBER
			MAIL DATE 05/18/2007	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/736,273	HOULE, MICHAEL EDWARD	
	Examiner	Art Unit	
	Miranda Le	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 30 January 2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-14 is/are pending in the application.
- 4a) Of the above claim(s) 10-14 is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-9 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This communication is responsive to Amendment, filed 01/30/07.

Claims 1-9 are pending in this application. Claims 1, 4, 6, 8 have been amended. This action is made Final.

2. The rejection of claim 1 under 35 U.S.C. §101 has been withdrawn in view of the amendment.

Election/Restrictions

3. Applicant's election with traverse of claims 1-9 is acknowledged.

Drawings

4. The drawings were received on 01/30/07. These drawings are acceptable.

Claim Objections

5. Claim 1 objected to because of the following informalities: Claim 1, line 17, "said inner-patch confidence values" should be changed to "said intra-patch confidence values".

Appropriate correction is required.

Claim Rejections - 35 USC § 103

6. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person

having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

7. Claims 1-5, 8, 9 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tang (U.S. Patent No. 6,636,849), in view of Kanno (U.S. Patent No. 7,007,019), and further in view of Gilmour (U.S. Patent No. 6,377,949).

As per claim 1, Tang teaches a computer system for generating data structures for information retrieval of documents stored in a database, said documents being stored as document-keyword vectors generated from a predetermined keyword list, and said document-keyword vectors forming nodes of a hierarchical structure imposed upon said documents (*i.e.*, *textual or byte-based searches, literature search based on lists of keywords, and vector and matrix based indexing and searching, Abstract*), said computer system comprising:

a neighborhood patch generation subsystem for generating groups of nodes having similarities as determined using a search structure (*i.e. a multigrid tree, col. 4, lines 39-54*), said neighborhood patch generation subsystem including a subsystem for generating a hierarchical structure upon said document-keyword vectors and a patch defining subsystem for creating patch relationships among said nodes (*i.e. multiple level of grids, col. 11, lines 18-27*) with respect to a metric distance between nodes (*i.e. L1 distance, col. 10, line 61 to col. 11, line 5*) (Figs. 5, 8A-B) (*col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

a query vector (*query point q, col. 4, line 55 to col. 5, line 3*) generation subsystem accepting query keywords (*i.e. keyword, col. 7, lines 32-41*), generating a corresponding query vector (*i.e. Once created, the multigrid tree can be searched to find exact or approximate or homologous matches for a search query. The multigrid tree can be searched to provide a solution to the following example. Suppose a multigrid search tree representing a set E in a metric space, and a query point in the metric space is provided, col. 11, line 64 to col. 12, line 31*);

intra-path confidence values (*i.e. $d(Sa, Sb)$, col. 16, lines 25-39; (i.e. a radius, col. 11, lines 6-18)*);

interpath confidence values (*i.e. $len(SaSb)$, col. 16, lines 25-39*);

a cluster estimation subsystem for generating cluster data of said document-keyword vectors (*i.e. a subset of grids, col. 11, lines 6-18*) using said similarities of patches (*col. 4, line 55 to col. 5, line 17, col. 10, line 61 to col. 11, line 27; col. 13, line 61 to col. 14, line 16*).

Tang does not explicitly teach:

a query vector generation subsystem accepting search conditions and query keywords, generating a corresponding query vector, and storing the generated query vector;

a confidence determination subsystem for computing intra-patch confidence values between patches and interpath confidence values;

the cluster estimation subsystem selects said patches depending on said intra-patch confidence values to represent clusters of said document keyword vectors, estimates the size of said patches and generates cluster data of document keyword vectors using similarities of the patches.

Kanno teaches:

a query vector (*i.e. query vector, col. 15, lines 34-40*) generation subsystem accepting search conditions (*i.e. Partial query condition, col. 15, lines 34-40*) and query keywords, generating a corresponding query vector (*i.e. Partial query condition calculation means 303 calculates a partial inner product lower limit value f as a lower limit value of an inner product of 37 types of 8-dimensional partial query vectors q with the partial vector corresponding to q by $f = \alpha \cdot |q|_{\text{sup.2}} / |Q|_{\text{sup.2}}$ with respect to partial spaces of 0 to 36 for the query vector Q obtained by the search condition input means 302, col. 15, lines 34-40*), and storing the generated query vector (*i.e. a vector database 101 stores 200,000 pieces of vector data constituted of two items of: a 296-dimensional unit real vector prepared from a newspaper article full text database of 200,000 collected newspaper articles and indicating characteristic of each newspaper article; and an identification number in a range of 1 to 200,000, and has a content as shown in FIGS. 12A and 12B, col. 5, lines 1-11*);

wherein the cluster estimation subsystem selects said patches depending on said intra-patch confidence values to represent clusters of said document keyword vectors, estimates the size of said patches (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [r.1, r.2])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*), and generates cluster data (*i.e. number of articles, col. 23, lines 6-29*) of document keyword vectors using similarities of the patches (*i.e. Constitution of Similar Vector Searching Apparatus, col. 21, line 64 to col. 22, lines 54*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang and Kanno at the time the invention was made to modify the system of Tang to include the limitations as taught by Kanno.

One of ordinary skill in the art would be motivated to make this combination in order to calculate an accumulated value of a partial square distance difference and used as a clue to the similarity search in view of Kanno, as doing so would give the added benefit of enabling a high-speed search to be possible with respect to the vector database even when the vector is of several hundreds of dimensions, as taught by Kanno (*col. 3, line 52 to col. 4 line 3*).

Tang, Kanno do not specifically teach a confidence determination subsystem for computing intra-patch confidence values between patches and interpath confidence values.

Gilmour teaches:

a confidence determination subsystem for computing intra-patch confidence values between patches and interpath confidence values (*i.e. confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang, Kanno and Gilmour at the time the invention was made to modify the system of Tang and Kanno to include a confidence determination subsystem for computing intra-patch confidence values between patches and interpath confidence values as taught by Gilmour.

One of ordinary skill in the art would be motivated to make this combination in order to assign a confidence level to a term within an electronic document in view of Gilmour (*col. 2, lines 19-28*), as doing so would give the added benefit of capturing knowledge automatically,

without excessive invasion or disruption of normal work patterns of participating users as taught by Gilmour (*col. 4, lines 44-57*).

As per claim 4, Tang teaches a method for generating data structures for information retrieval of documents stored in a database, said documents being stored as document-keyword vectors generated from a predetermined keyword list, and said document-keyword vectors forming nodes of a hierarchical structure imposed upon said documents (*i.e., textual or byte-based searches, literature search based on lists of keywords, and vector and matrix based indexing and searching, Abstract*), said method comprising the steps of:

generating a hierarchical structure (*i.e. multigrid tree, col. 4, lines 39-54*) upon said document-keyword vectors and storing hierarchy data in an adequate storage area (*Figs. 5, 8A-B*) (*col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

accepting query keywords (*i.e. keyword, col. 7, lines 32-41*), generating (*i.e. Once created, the multigrid tree can be searched to find exact or approximate or homologous matches for a search query. The multigrid tree can be searched to provide a solution to the following example. Suppose a multigrid search tree representing a set E in a metric space, and a query point in the metric space is provided, col. 11, line 64 to col. 12, line 31*) a corresponding query vector (*query point q, col. 4, line 55 to col. 5, line 3*);

generating neighborhood patches of nodes having similarities as determined using levels of the hierarchical structure (*i.e. multiple level of grids, col. 11, lines 18-27*), and storing said patches in an adequate storage area (*Figs. 5, 8A-B; col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

intra-path confidence values (*i.e. $d(S_a, S_b)$, col. 16, lines 25-39; (i.e. a radius, col. 11, lines 6-18);*

interpath confidence values (*i.e. $len(S_a S_b)$, col. 16, lines 25-39);*

Tang does not fairly teach:

accepting search conditions and query keywords, generating a corresponding query vector, and storing the generated query vector;

determining intra-patch confidence values between patches and interpatch confidence values;

invoking said hierarchy data and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area;

estimating the sizes of said patches, and generating cluster data of document keyword vector using similarities of the patches, selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors.

Kanno teaches:

accepting search conditions and query keywords (*i.e. Partial query condition, col. 15, lines 34-40*), generating a corresponding query vector (*i.e. Partial query condition calculation means 303 calculates a partial inner product lower limit value f as a lower limit value of an inner product of 37 types of 8-dimensional partial query vectors q with the partial vector corresponding to q by $f = \frac{\alpha \cdot |q| \cdot \sup 2}{|Q| \cdot \sup 2}$ with respect to partial spaces of 0 to 36 for the query vector Q obtained by the search condition input means 302, col. 15, lines 34-40*), and

Art Unit: 2167

storing the generated query vector (*i.e. a vector database 101 stores 200,000 pieces of vector data constituted of two items of: a 296-dimensional unit real vector prepared from a newspaper article full text database of 200,000 collected newspaper articles and indicating characteristic of each newspaper article; and an identification number in a range of 1 to 200,000, and has a content as shown in FIGS. 12A and 12B, col. 5, lines 1-11*);

estimating the sizes of said patches (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [l, r])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*), and generating cluster data of document keyword vector using similarities of the patches (*i.e. Constitution of Similar Vector Searching Apparatus, col. 21, line 64 to col. 22, lines 54*), selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [l, r])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang and Kanno at the time the invention was made to modify the system of Tang to include the limitations as taught by Kanno.

One of ordinary skill in the art would be motivated to make this combination in order to calculate an accumulated value of a partial square distance difference and used as a clue to the similarity search in view of Kanno, as doing so would give the added benefit of enabling a high-speed search to be possible with respect to the vector database even when the vector is of several hundreds of dimensions, as taught by Kanno (*col. 3, line 52 to col. 4 line 3*).

Tang, Kanno do not expressly teach determining intra-patch confidence values between patches and interpatch confidence values;

invoking said hierarchy data and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage.

Gilmour teaches:

determining intra-patch confidence values between patches and interpatch confidence values (*i.e. confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62*);

invoking said hierarchy data (*i.e. a hierarchical structure of classes, col. 9, lines 16-44*) and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area (*i.e. confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang, Kanno and Gilmour at the time the invention was made to modify the system of Tang and Kanno to include the limitations as taught by Gilmour.

One of ordinary skill in the art would be motivated to make this combination in order to assign a confidence level to a term within an electronic document in view of Gilmour (*col. 2, lines 19-28*), as doing so would give the added benefit of capturing knowledge automatically, without excessive invasion or disruption of normal work patterns of participating users as taught by Gilmour (*col. 4, lines 44-57*).

As per claim 8, Tang teaches a computer readable medium storing a program for making a computer system execute a method for generating data structures for information retrieval of documents stored in a database, said documents being stored as document-keyword vectors generated from a predetermined keyword list, and said document-keyword vectors forming nodes of a hierarchical structure imposed upon said documents (*i.e.* , *textual or byte-based searches, literature search based on lists of keywords, and vector and matrix based indexing and searching, Abstract*), said program making said computer system execute the steps of:

accepting query keywords (*i.e. keyword, col. 7, lines 32-41*), generating (*i.e. Once created, the multigrid tree can be searched to find exact or approximate or homologous matches for a search query. The multigrid tree can be searched to provide a solution to the following example. Suppose a multigrid search tree representing a set E in a metric space, and a query point in the metric space is provided, col. 11, line 64 to col. 12, line 31*) a corresponding query vector (*query point q, col. 4, line 55 to col. 5, line 3*)

generating a hierarchical structure (*i.e. multigrid tree, col. 4, lines 39-54*) upon said document-keyword vectors and storing hierarchy data in an adequate storage area (*Figs. 5, 8A-B*) (*col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

generating neighborhood patches consisting of nodes having similarities as determined using levels of the hierarchical structure (*i.e. multiple level of grids, col. 11, lines 18-27*), and storing said patches list in an adequate storage area (*Figs. 5, 8A-B; col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

Tang does not specifically teach:

accepting search conditions and query keywords, generating a corresponding query vector, and storing the generated query vector;

invoking said hierarchy data and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area;

selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors.

Kanno teaches:

accepting search conditions and query keywords (*i.e. Partial query condition, col. 15, lines 34-40*), generating a corresponding query vector (*i.e. Partial query condition calculation means 303 calculates a partial inner product lower limit value f as a lower limit value of an inner product of 37 types of 8-dimensional partial query vectors q with the partial vector corresponding to q by $f = \frac{\alpha \cdot |q| \cdot \sup.2}{|Q| \cdot \sup.2}$ with respect to partial spaces of 0 to 36 for the query vector Q obtained by the search condition input means 302, col. 15, lines 34-40*), and storing the generated query vector (*i.e. a vector database 101 stores 200,000 pieces of vector data constituted of two items of: a 296-dimensional unit real vector prepared from a newspaper article full text database of 200,000 collected newspaper articles and indicating characteristic of each newspaper article; and an identification number in a range of 1 to 200,000, and has a content as shown in FIGS. 12A and 12B, col. 5, lines 1-11*);

selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors (*i.e. Search*

object range generation means 404 enumerates all sets (d, c, [1, r.2]) of the region number d for specifying a region including a partial vector, col. 22, lines 30-39).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang and Kanno at the time the invention was made to modify the system of Tang to include the limitations as taught by Kanno.

One of ordinary skill in the art would be motivated to make this combination in order to calculate an accumulated value of a partial square distance difference and used as a clue to the similarity search in view of Kanno, as doing so would give the added benefit of enabling a high-speed search to be possible with respect to the vector database even when the vector is of several hundreds of dimensions, as taught by Kanno (*col. 3, line 52 to col. 4 line 3*).

Tang, Kanno do not specifically teach:

invoking said hierarchy data and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area.

Gilmour teaches:

invoking said hierarchy data (*i.e. a hierarchical structure of classes, col. 9, lines 16-44*) and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area (*i.e. confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang, Kanno and Gilmour at the time the invention was made to modify the system of Tang and Kanno to include the limitations as taught by Gilmour.

One of ordinary skill in the art would be motivated to make this combination in order to assign a confidence level to a term within an electronic document in view of Gilmour (*col. 2, lines 19-28*), as doing so would give the added benefit of capturing knowledge automatically, without excessive invasion or disruption of normal work patterns of participating users as taught by Gilmour (*col. 4, lines 44-57*).

As per claim 2, Kanno teaches the computer system of claim 1, wherein said cluster estimation subsystem selects said patches depending on said inter-patch confidence values to represent clusters of said document-keyword vectors (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [.1, r.2])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*).

As per claim 3, Kanno teaches the computer system of claim 1, wherein said cluster estimation subsystem estimates sizes of said clusters depending on said intra-patch confidence values (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [.1, r.2])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*).

As per claim 5, Kanno teaches the method according to claim 4 further comprising the step of estimating sizes of said clusters depending on said intra-patch confidence values (*i.e.*

Search object range generation means 404 enumerates all sets (d, c, [.1, r.2]) of the region number d for specifying a region including a partial vector, col. 22, lines 30-39).

As per claim 9, Kanno teaches the method according to claim 8, further comprising the step of estimating sizes of said clusters depending on said intra-patch confidence values (*i.e.* *Search object range generation means 404 enumerates all sets (d, c, [.1, r.2]) of the region number d for specifying a region including a partial vector, col. 22, lines 30-39).*

8. Claims 6, 7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Tang (U.S. Patent No. 6,636,849), in view of Gilmour (U.S. Patent No. 6,377,949), and further in view of Kanno (U.S. Patent No. 7,007,019).

As per claim 6, Tang teaches a program for making a computer system execute a method for generating data structures for information retrieval of documents stored in a database, said documents being stored as document-keyword vectors generated from a predetermined keyword list, and said document-keyword vectors forming nodes of a hierarchical structure introduced into said documents (*i.e. textual or byte-based searches, literature search based on lists of keywords, and vector and matrix based indexing and searching, Abstract*), said program making said computer system execute the steps of:

generating a hierarchical structure (*i.e. multigrid tree, col. 4, lines 39-54*) upon said document-keyword vectors and storing hierarchy data in an adequate storage area (*Figs. 5, 8A-B*) (*col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

generating neighborhood patches consisting of nodes having similarities as determined using levels of the hierarchical structure (*i.e. multiple level of grids, col. 11, lines 18-27*), and storing said patches in an adequate storage area (*Figs. 5, 8A-B; col. 4, line 55 to col. 5, line 17, col. 13, line 61 to col. 14, line 16*);

intra-path confidence values (*i.e. $d(S_a, S_b)$, col. 16, lines 25-39; (i.e. a radius, col. 11, lines 6-18)*);

interpath confidence values (*i.e. $len(S_a S_b)$, col. 16, lines 25-39*);

Tang does not fairly teach:

determining intra-patch confidence values between patches and interpatch confidence values;

invoking said hierarchy data and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area;

selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors.

Gilmour teaches:

determining intra-patch confidence values between patches and interpatch confidence values (*i.e. confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62*).

invoking said hierarchy data (*i.e. a hierarchical structure of classes, col. 9, lines 16-44*) and said patches to compute inter-patch confidence values between said patches and intra-patch confidence values, and storing said values as corresponding lists in an adequate storage area (*i.e.*

confidence level values for the term is then determined based on the summed adjusted counts and the term weight, col. 16, lines 32-62).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang, and Gilmour at the time the invention was made to modify the system of Tang to include the limitations as taught by Gilmour.

One of ordinary skill in the art would be motivated to make this combination in order to assign a confidence level to a term within an electronic document in view of Gilmour (*col. 2, lines 19-28*), as doing so would give the added benefit of capturing knowledge automatically, without excessive invasion or disruption of normal work patterns of participating users as taught by Gilmour (*col. 4, lines 44-57*).

Tang, Gilmour do not explicitly teach:

selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors.

Kanno teaches:

selecting said patches depending on said inter-patch confidence values and said intra-patch confidence values to represent clusters of said document-keyword vectors (*i.e. Search object range generation means 404 enumerates all sets $(d, c, [1, r.2])$ of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*).

It would have been obvious to one of ordinary skill of the art having the teaching of Tang, Gilmour and Kanno at the time the invention was made to modify the system of Tang and Gilmour to include the limitations as taught by Kanno.

One of ordinary skill in the art would be motivated to make this combination in order to calculate an accumulated value of a partial square distance difference and used as a clue to the similarity search in view of Kanno, as doing so would give the added benefit of enabling a high-speed search to be possible with respect to the vector database even when the vector is of several hundreds of dimensions, as taught by Kanno (*col. 3, line 52 to col. 4 line 3*).

As per claim 7, Kanno teaches the method according to claim 6, further comprising the step of estimating sizes of said clusters depending on said intra-patch confidence values (*i.e. Search object range generation means 404 enumerates all sets (d, c, [1, r.2]) of the region number d for specifying a region including a partial vector, col. 22, lines 30-39*).

Response to Arguments

9. Applicant's arguments with respect to claims 1-11 have been considered but are moot in view of the new ground(s) of rejection.

Conclusion

9. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after

the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Miranda Le whose telephone number is (571) 272-4112. The examiner can normally be reached on Monday through Friday from 8:30 AM to 5:00 PM.


If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John R. Cottingham, can be reached on (571) 272-7079. The fax number to this Art Unit is 571-273-8300.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (703) 305-3900.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



Miranda Le
May 11, 2007



JOHN COTTINGHAM
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100